



## 【特許請求の範囲】

【請求項1】計算機内に記録した対象物体の3次元形状モデルから対象物体をカメラで撮影した入力画像上での見え方を算出するグラフィックスシステムが、その過程で提供する対象シーンの奥行き画像を利用し、画像上に観測される遮蔽輪郭線に対応するモデル上の遮蔽輪郭生成点の3次元座標を、対象物体の複雑さによらず、算出する方法。

【請求項2】計算機内に記録した対象物体の3次元形状モデルから対象物体をカメラで撮影した入力画像上での見え方を算出するグラフィックスシステムが、その過程で提供する対象シーンの奥行き画像を利用して得た遮蔽輪郭生成点の3次元座標と観測画像上の2次元座標との対応関係から対象物体の位置・姿勢を算出する方法。

【請求項3】計算機内に記録した環境の3次元形状モデルから、その環境内のある位置・姿勢に置かれたカメラで撮影したその見え方を算出するグラフィックスシステムが、その過程で提供する対象シーンの奥行き画像を利用し、環境モデル上の遮蔽輪郭生成点の3次元座標と観測画像上の2次元座標との対応関係から撮影しているカメラの位置・姿勢を相対的に算出する方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、観測画像から対象物体や観測カメラの位置・姿勢を算出する方法に関する。この技術が実現すると、ロボットによる作業の自動化などが大きく促進される。

## 【0002】

【従来の技術】ロボットが室内を移動し作業を行う場合、作業対象物体の位置・姿勢、及び、全体の環境におけるロボットの位置・姿勢を正確に把握する必要がある。こうした位置・姿勢情報は、ロボットの目である、ロボットに搭載されたカメラにより入力される画像情報と、ロボット内に持つ環境や対象物体の3次元形状モデル間点対応により、次のように得ることができる。

## 【0003】

## 【数1】

3次元モデル点 :  $\mathbf{X}_i(X_i, Y_i, Z_i)^\top$

観測点 :  $\mathbf{x}_i(x_i, y_i, f)^\top$

のm個のペアが得られているとする。この対応を用いて

## 【数2】

$$\begin{aligned} x_i &= f X'_i / Z'_i, \quad y_i = f Y'_i / Z'_i \\ \mathbf{X}'_i &= \mathbf{R} \mathbf{X}_i + \mathbf{T} \end{aligned}$$

の関係式を満たす3次元モデルの剛体運動、

## 【数3】

3×3回転行列 :  $\mathbf{R}$

平行移動ベクトル :  $\mathbf{T}$

を算出することにより、観測状態の位置・姿勢が求まる。

## 【0004】

【発明が解決しようとする課題】対象が特徴的なマーク点や角点を持つ場合には、観測画像からの特徴点抽出も、モデル上の特徴点との対応付けも比較的容易である。しかし、図1のように、観測画像上に観測される特徴が遮蔽輪郭の場合、この遮蔽輪郭特徴に対応するモデル表面上の点は視点によって変化するため、2次元画像上に観測される点と3次元モデル上の対応点を知るのが非常に難しい。結果的にこうした観測遮蔽輪郭から対象や環境の位置・姿勢を求めることがほとんど不可能であるという問題があった。

## 【0005】

【課題を解決するための手段】対象の大まかな位置・姿勢は既知である条件のもとで、上記課題を達成するために、本発明では、計算機内部に記録した対象物体の3次元形状モデルから、この大まかな初期位置・姿勢における見え方をグラフィックスシステムで算出し、この過程で提供される対象シーンの奥行き画像を利用し、画像上に観測される遮蔽輪郭に対する、モデル上の対応候補点の3次元座標を高速に算出する方法を提供する。(請求項1)

【0006】さらに、こうして得た遮蔽輪郭特徴の2D-3D対応ペアをもとに、初期に与えられた大まかなモデルの位置・姿勢を繰り返し処理により正しい状態に更新していく、最終的には正しい対象の位置・姿勢を得る方法を提供する。(請求項2)

さらに、これを計算機内部にカメラを取り巻く環境の3次元形状モデルが記録されている場合に拡張し、観測するカメラの位置・姿勢を算出する方法を提供する。(請求項3)

【0007】本願発明の方法によると、対象形状を限定することなく、例え複数の物体により構成される複雑な環境モデルの場合にも各個体間の遮蔽や自己遮蔽などがすべて自動的に考慮されるその性格により、同様に処理が実現される。

## 【0008】

【発明の実施の形態】図2と図3を用い、遮蔽輪郭を用いて対象物体の位置・姿勢を算出する処理を説明する。

ステップS1 : カメラにより画像を入力する。

ステップS2 : 画像から明るさが急変するエッジを抽出する。

ステップS3 : モデル上の遮蔽輪郭生成点の3次元座標の算出。

【0009】図3aは計算機内部に記録されている対象物体の幾何モデルである。今、この対象の大まかな位置・姿勢が既知であれば、OpenGL (Silicon Graphics社の登録商標) に代表されるグラフィックスシステムにより、その見え方の予測画像、図3cを高速に算出するこ

とが可能である。この際、グラフィックスシステムはこの画像の算出のため、同時に図 3d のような奥行き画像も算出する。これは、各画素の示す視線方向に向かって、どれだけ離れた位置にモデルが存在するかの距離を画素ごとに記録したものである。図 3d の例では、この距離が近いほど明るく表示してある。この画像からエッジ抽出処理を行うことにより、図 3e 中のグレイ線のように奥行きが大きく変わる画素が抽出され、その境界隣接画素のうち、手前側（距離値が小さいもの）をとることにより、遮蔽輪郭に対応するモデル表面の 3 次元座標値が得られる。ただし、この奥行き画像を作成するためには用いた、対象物体の位置・姿勢の初期値は実際の値とは異なっているため、多少の誤差を含む。多少の誤差を含むが、これら遮蔽輪郭生成 3D 点は以下に続く処理によって、対象の位置・姿勢認識に利用することが可能である。

【0010】ステップ S4：観測エッジと対応する 3 次元モデル代表点の算出。

上記手順により得られた遮蔽輪郭生成 3 次元点を投影画  
On rotation:

$$\min_{\mathbf{R}} \sum_{k=1}^n \sum_{i=1}^{m(c_k)-1} \sum_{j=i+1}^{m(c_k)} (F_{c_k}(\mathbf{x}_i^{c_k} \times \mathbf{x}_j^{c_k}) \cdot \mathbf{R}(\mathbf{X}_i - \mathbf{X}_j))^2 \quad (1)$$

【数 5】

像上で等間隔にサンプリングすることにより、図 3e に白い点で示すような、代表点を求める。さらに、予測画像を利用して、このうち画像上で明確なエッジとして検出される見通しのものだけを残す（図 3f）。

【0011】ステップ S5：3 次元モデル点と観測エッジ点との対応付。

テリトリベース 3D-2D レジストレーション手法「喜多 泰代、 Dale L. Wilson、 J. Alison Noble: '3 次元血管モデルの X 線画像への実時間位置合わせ手法'、電子情報通信学会論文誌 (D-II), Vol. J83-DII, No. 1, pp. 254-262 (2000):」の部分では、観測画像上に投影したこれらモデル点と観測エッジ点で、画像上の 2 次元距離の近さを評価基準として、点対応ペアを作成する。

【0012】ステップ S6：対象物体の位置・姿勢を算出。

これらの点対応ペアを、n 個のカメラの観測画像から対象の位置・姿勢を算出する一般式、

【数 4】

On translation:

$$\begin{bmatrix}
 f^{c_1} \lambda_x^{c_1} - x_1^{c_1} \lambda_z^{c_1} & f^{c_1} \mu_x^{c_1} - x_1^{c_1} \mu_z^{c_1} & f^{c_1} \nu_x^{c_1} - x_1^{c_1} \nu_z^{c_1} \\
 f^{c_1} \lambda_y^{c_1} - y_1^{c_1} \lambda_z^{c_1} & f^{c_1} \mu_y^{c_1} - y_1^{c_1} \mu_z^{c_1} & f^{c_1} \nu_y^{c_1} - y_1^{c_1} \nu_z^{c_1} \\
 \vdots & \vdots & \vdots \\
 f^{c_1} \lambda_x^{c_1} - x_{m(c_1)}^{c_1} \lambda_z^{c_1} f^{c_1} \mu_x^{c_1} - x_{m(c_1)}^{c_1} \mu_z^{c_1} f^{c_1} \nu_x^{c_1} - x_{m(c_1)}^{c_1} \nu_z^{c_1} \\
 f^{c_1} \lambda_y^{c_1} - y_{m(c_1)}^{c_1} \lambda_z^{c_1} f^{c_1} \mu_y^{c_1} - y_{m(c_1)}^{c_1} \mu_z^{c_1} f^{c_1} \nu_y^{c_1} - y_{m(c_1)}^{c_1} \nu_z^{c_1} \\
 f^{c_2} \lambda_x^{c_2} - x_1^{c_2} \lambda_z^{c_2} & f^{c_2} \mu_x^{c_2} - x_1^{c_2} \mu_z^{c_2} & f^{c_2} \nu_x^{c_2} - x_1^{c_2} \nu_z^{c_2} \\
 f^{c_2} \lambda_y^{c_2} - y_1^{c_2} \lambda_z^{c_2} & f^{c_2} \mu_y^{c_2} - y_1^{c_2} \mu_z^{c_2} & f^{c_2} \nu_y^{c_2} - y_1^{c_2} \nu_z^{c_2} \\
 \vdots & \vdots & \vdots
 \end{bmatrix} \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \\
 = \begin{bmatrix}
 x_1^{c_1} Z_1'^{c_1} - f^{c_1} X_1'^{c_1} \\
 y_1^{c_1} Z_1'^{c_1} - f^{c_1} Y_1'^{c_1} \\
 \vdots \\
 x_{m(c_1)}^{c_1} Z_{m(c_1)}'^{c_1} - f^{c_1} X_{m(c_1)}'^{c_1} \\
 y_{m(c_1)}^{c_1} Z_{m(c_1)}'^{c_1} - f^{c_1} Y_{m(c_1)}'^{c_1} \\
 x_1^{c_2} Z_1'^{c_2} - f^{c_2} X_1'^{c_2} \\
 y_1^{c_2} Z_1'^{c_2} - f^{c_2} Y_1'^{c_2} \\
 \vdots
 \end{bmatrix} \quad (2)$$

に  $n = 1$  の条件で代入することにより、この対応ペアを  
最小自乗誤差の意味において満す、対象の位置・姿勢が

算出される。ここで、

$(\lambda_{x(y,z)}^{c_k}, \mu_{x(y,z)}^{c_k}, \nu_{x(y,z)}^{c_k})$  : k 番目のカメラ座標系の  $x^{c_k}(y^{c_k}, z^{c_k})$  軸の  
方向余弦

 $f^{c_k}$  $F_{c_k}$  $m(c_k)$ 

30

【0013】

【数6】

: k 番目のカメラの焦点距離

: k 番目のカメラ座標系の値をワールド

座標系の値に変換する関数

: k 番目のカメラ座標系の観測点と

3 次元モデル点の対応点ペア数

である。ただし、画像上の距離の近さに基づき定めたペアが完全に正しいペアではないこと、この一般式が回転行列の線形化による近似誤差を含んでいること、遮蔽輪郭生成の3次元座標が多少誤差を含んでいることから、一回の演算で正しい対象の位置・姿勢は得られない。そこで、新しい位置・姿勢の値を用いて、ステップ3からS6までの処理を対象の位置・姿勢が収束するまで繰り返すことにより、最終的に正しい位置・姿勢を算出する。

【0014】観測カメラの位置・姿勢検出は、計算機内部にカメラを取り巻く環境の形状モデルを記録し、これ

40 を対象形状モデルとして上記と同様に対象物体の新しい位置・姿勢を算出し、これと相対的にカメラの位置・姿勢を変化させ、ステップ3に戻る、という繰り返し処理によって、得られる。

【0015】カメラを取り巻く環境形状モデルを計算機に与え、環境におけるカメラの位置・姿勢を求める実施例を以下に示す。図4aは、実際の観測画像である。図4bはこのエッジ画像 (Canny エッジオペレータ「一般的に用いられている画像処理手法」: J. Canny: 'A Computational Approach to Edge Detection', IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 8, N

o. 6, pp. 679-698. を使用。) で、各エッジ点近傍で濃淡値の最大変化方向を 8 方向に濃淡差をつけて示してある。図 4 c に OpenGL で作成した円柱 5 本からなる部分モデルの側面図と上から見た図を示す。上面図中の白丸はカメラ位置、そこから出ている白線は視線方向を示す。側面図は、観測画像を得たカメラの初期位置・姿勢を大まかにマニュアルで計測し、その状態における予測画像となっている。図 4 d は、これを観測画像に重ねて表示したもので、初期位置・姿勢計測誤差のため実際に各パイプが観測される位置からずれている。見易いようにパイプを黒で表示している。図 4 e に、予測奥行き画像のエッジ上に求めた 3 次元モデル点(白い点)を側面図、上面図に示す。ただし、予測濃淡画像情報を用いて明るさ変化が少なく観測されにくい点は省かれるため、同色のパイプ同士が重なっているところは選択されていない。観測画像のエッジ点と同様、予測投影画像上のモデル点近傍の濃淡値最大変化方向を 8 方向に分類し各モデル点の特性として持つ。図 4 f に、この特性を濃淡差で表し、その投影位置を観測エッジ画像に重ねて表示する。

【0016】カメラ角度のわずかなずれが、画面上では大きな平行移動を生むので、第1回目の初期対応時だけはこれを考慮する。具体的に、投影モデル点を画面上で 2 次元的に平行移動し、モデル点が同色のエッジ位置に最もよく重なる位置を探索する。図 4 g が移動後の位置である。これはあくまで最短距離評価基準でモデル点の対応ペアを見つけるために必要なので、実際に 3 次元モデルを移動する必要はない。この位置においてテリトリベース 3D-2D マッチングを行い、得られた対応付が同図中に白線で示されている。式(1), (2) を用いて得られた位置・姿勢変化量を相対的にカメラの位置・姿勢に反映し、その新しい位置において同様の処理を繰り返す(2 回目以降はすでに画面上での大きなずれはなくなっている。画像上の平行移動は不要)。この例では、24 回繰り返し後に収束条件を満し、カメラ位置を(66.8, -13.2, -80.7) (mm) 平行移動修正し、軸(-0.34, 0.37, -0.86) 回りに 5.7 度回転修正して、図 4 h, i のような予測画像と観測画像の重なりを得た。モデルの投影位置が観測画像と

一致しているということは、カメラの位置・姿勢がより正しい値に修正されたことを示している。処理時間は、Pentium (登録商標) 11 (333MHz) マシン上で、画像上の初期平行移動に 2.1sec、3D モデル点を選出し、その対応観測特徴点を求め、これに基づきカメラの位置・姿勢を変更する 1 ループに約 200 msec かかり、トータルで 6-8 秒であった。

#### 【0017】

【発明の効果】従来、観測画像上の遮蔽輪郭を対象物体の位置・姿勢検出、また、相対的に観測カメラの位置・姿勢検出に利用することは非常に単純な対象物以外では不可能であった。本発明の方法は、あらかじめ、大まかな位置・姿勢が与えられる場合に、これを可能とした。これにより、特徴的な角点やマーク点などがない対象物体の位置・姿勢検出、また特徴的な角点やマーク点などが存在せず、複雑な曲面で構成される環境中のカメラの位置・姿勢検出が可能となった。

#### 【図面の簡単な説明】

【図 1】遮蔽輪郭が特徴エッジの大半を占める環境例を示す図である。

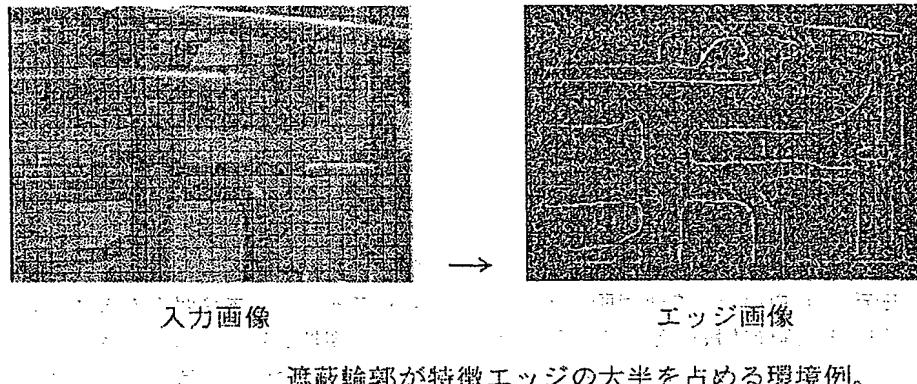
【図 2】遮蔽輪郭特徴を用いた対象物体の位置・姿勢検出の流れ図である。

【図 3】遮蔽輪郭特徴を用いた対象物体の位置・姿勢検出手法の説明図である。

【図 4】本発明の実施を例示する図である。

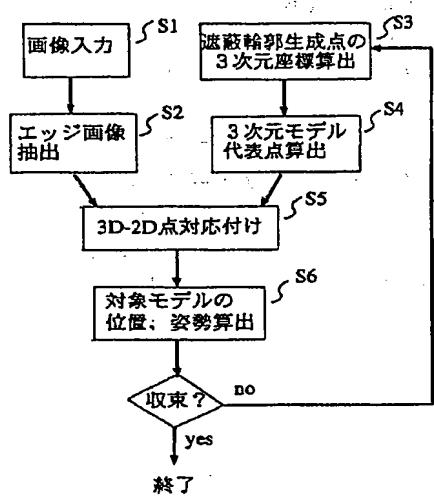
- (a) 観測画像;
- (b) 観測画像のエッジ(濃淡値の最大変化方向別に色分け);
- (c) 対象の 3 次元モデルの側面図および上面図;
- (d) カメラの初期推定位置・姿勢におけるモデル投影;
- (e) 3 次元モデル点(白い点)の正面図、側面図;
- (f) 3 次元モデル点の観測エッジ画像での投影位置;
- (g) 初期平行移動後の 3 次元投影モデル点;
- (h) 収束後のカメラ位置・姿勢におけるモデル点投影(エッジ画像上);
- (i) 収束後のカメラ位置・姿勢におけるモデル投影。

【図1】



遮蔽輪郭が特徴エッジの大半を占める環境例。

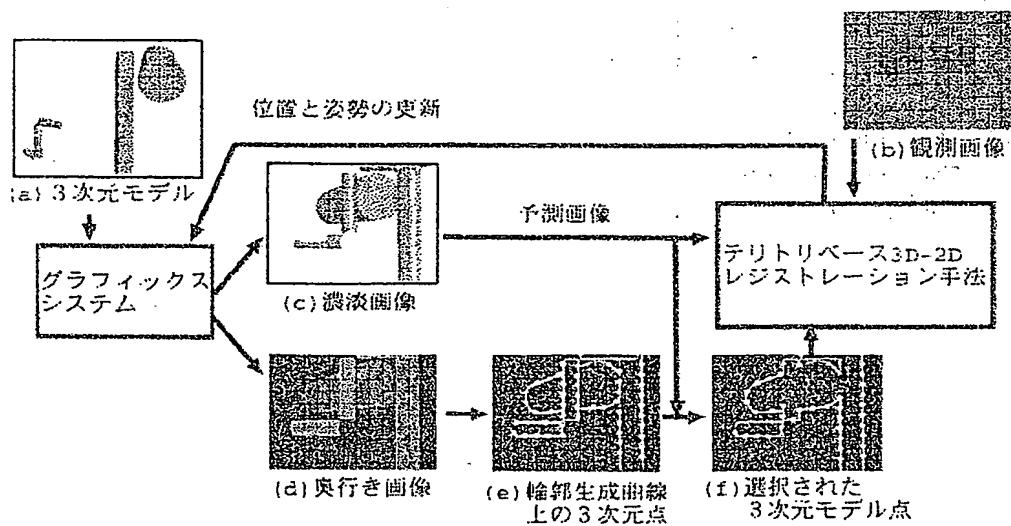
【図2】



遮蔽輪郭特徴を用いた対象物体の位置、姿勢検出の流れ図

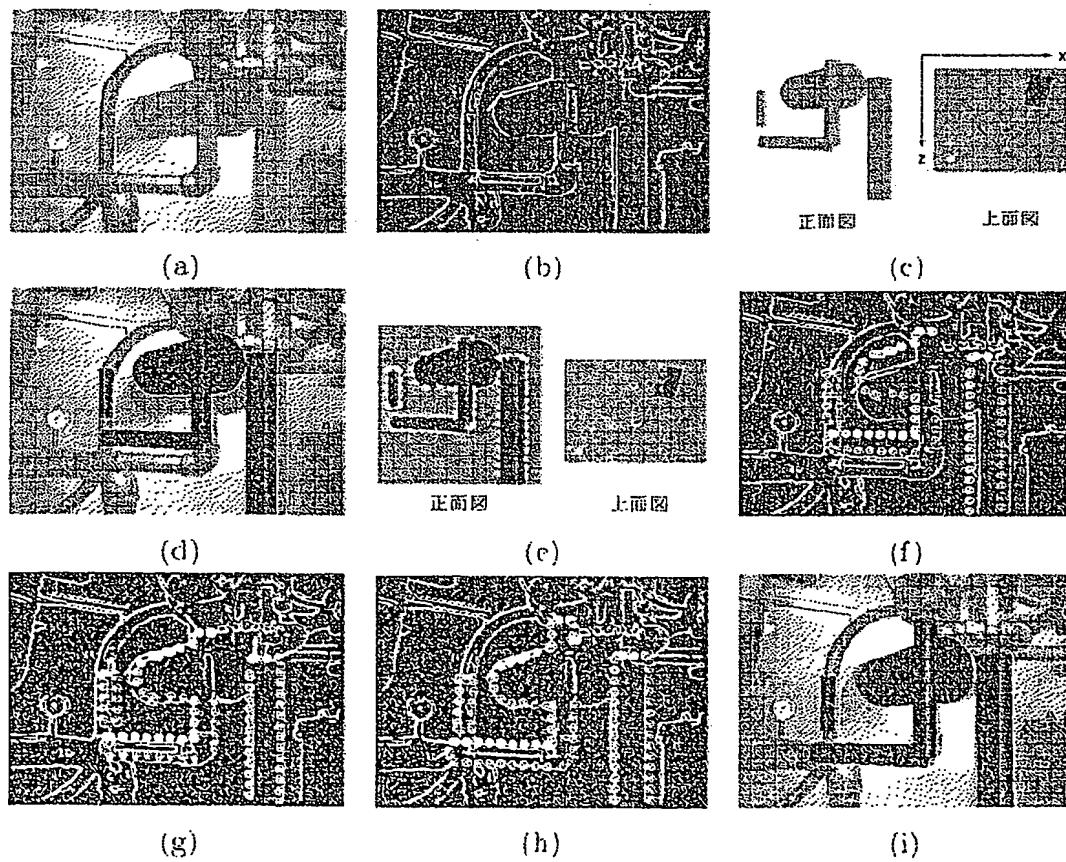
BEST AVAILABLE COPY

【図3】



遮蔽輪郭特徴を用いた対象物体の位置・姿勢検出手法の説明図

【図4】



## フロントページの続き

F ターム(参考) 3F059 BB04 DA02 DA08 DB09  
5B057 AA05 BA02 CA12 CA16 CB12  
CB16 CB19 DA07 DB02 DC16  
DC36  
5C054 AA01 EA05 FC14 FC15 FD02  
FD03 FE13 HA02 HA04 HA05  
5L096 BA05 CA02 FA06 GA32 JA11